STAT 400
UIUC

Hypergeometric and Multinomial Dist.

Stepanov
Dalpiaz
Nguyen

Example 1:

Suppose that on Halloween 6 children come to a house to get treats. A bag contains 8 plain chocolate bars and 7 nut bars. Each child reaches into the bag and randomly selects 1 candy bar. Let X denote the number of nut bars selected.

a) Is the Binomial model appropriate for this problem?

b) Find the probability that exactly 2 nut bars were selected.

Hypergeometric Distribution:

N = population size,

S = number of "successes" in the population,

$N - S$ = number of "failures" in the population,

n = sample size.

X = number of "successes" in the sample when sampling is done without replacement.

Then

$$P(X = x) = \frac{\binom{S}{x} \cdot \binom{N-S}{n-x}}{\binom{N}{n}} = \frac{{}_S C_x \cdot {}_{N-S} C_{n-x}}{{}_N C_n}$$

OR

$$P(X = x) = \binom{n}{x} \cdot \left[\frac{S}{N} \cdot \frac{S-1}{N-1} \cdot \ldots \cdot \frac{S-x+1}{N-x+1} \right] \cdot \left[\frac{N-S}{N-x} \cdot \frac{N-S-1}{N-x-1} \cdot \ldots \cdot \frac{N-S-(n-x)+1}{N-n+1} \right]$$

$\max(0, n + S - N) \le x \le \min(n, S)$.

R: `phyper(q, m, n, k)` `dhyper(x, m, n, k)`

Example 2:

A jar has N marbles, S of them are orange and $N - S$ are blue. Suppose 3 marbles are selected. Find the probability that there are 2 orange marbles in the sample, if the selection is done …

with replacement | without replacement

a) $N = 10,\ S = 4;$

b) $N = 100,\ S = 40;$

c) $N = 1{,}000,\ S = 400;$

	Binomial	**Hypergeometric**
	with replacement	without replacement
Probability	$P(X = x) = \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x}$	$P(X = x) = \dfrac{\binom{S}{x} \cdot \binom{N-S}{n-x}}{\binom{N}{n}}$
Expected Value	$E(X) = n \cdot p$	$E(X) = n \cdot \dfrac{S}{N}$
Variance	$Var(X) = n \cdot p \cdot (1-p)$	$Var(X) = n \cdot \dfrac{S}{N} \cdot \left(1 - \dfrac{S}{N}\right) \cdot \dfrac{N-n}{N-1}$

If the population size is large (**compared to the sample size**) Binomial Distribution can be used regardless of whether sampling is with or without replacement.

Multinomial Distribution:

- The number of trials, n, is fixed.
- Each trial has k possible outcomes, with probabilities $p_1, p_2, \ldots, p_k$, respectively.

  $(p_1 + p_2 + \ldots + p_k = 1)$
- The trials are independent.
- $X_1, X_2, \ldots, X_k$ represent the number of times outcome 1, outcome 2, … , outcome k occur, respectively. $(X_1 + X_2 + \ldots + X_k = n)$

Then

$$P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2! \ldots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k},$$

$$x_1 + x_2 + \ldots + x_k = n.$$

R: `dmultinom(x, prob)`

<u>Example 3</u>:
A particular brand of candy-coated chocolate comes in six different colors. Suppose 30% of all pieces are brown, 20% are blue, 15% are red, 15% are yellow, 10% are green, and 10% are orange. Thirty pieces are selected at random.

a) What is the probability that 10 are brown, 8 are blue, 7 are red, 3 are yellow, 2 are green, and none are orange?

b) What is the probability that 10 are brown, 8 are blue, and 12 are of other colors?

Example 4:

When Stéphane plays chess against his favorite computer program, he wins with probability 0.60, loses with probability 0.10, and 30% of the games result is a draw. Assume independence.

Now, assume Stéphane plays 12 games.

a) Find the probability that he wins 5 games, loses 3 games, and draws 4 games.

b) Find the probability that he wins 7 games, and draws 5 games.

c) Find the probability that Stéphane wins at least 8 games.